

# Weak Fourier-Schur sampling, the hidden subgroup problem, and the quantum collision problem

Andrew M. Childs\*

Institute for Quantum Information  
California Institute of Technology  
Pasadena, CA 91125, USA

Aram W. Harrow†

Department of Computer Science  
University of Bristol  
Bristol, BS8 1UB, UK

Paweł Wocjan‡

School of Electrical Engineering and Computer Science  
University of Central Florida  
Orlando, FL 32816, USA

## Abstract

Schur duality decomposes many copies of a quantum state into subspaces labeled by partitions, a decomposition with applications throughout quantum information theory. Here we consider applying Schur duality to the problem of distinguishing coset states in the standard approach to the hidden subgroup problem. We observe that simply measuring the partition (a procedure we call *weak Schur sampling*) provides very little information about the hidden subgroup. Furthermore, we show that under quite general assumptions, even a combination of weak Fourier sampling and weak Schur sampling fails to identify the hidden subgroup. We also prove tight bounds on how many coset states are required to solve the hidden subgroup problem by weak Schur sampling, and we relate this question to a quantum version of the collision problem.

## 1 Introduction

The hidden subgroup problem (HSP) is a central challenge for quantum computation. On the one hand, many of the known fast quantum algorithms are based on the efficient solution of the abelian HSP [26, 27, 46, 50]. On the other hand, the *nonabelian* hidden subgroup problem has potential algorithmic applications: in particular, the graph isomorphism problem can be reduced to the HSP in the symmetric group [10, 17], and the shortest lattice vector problem can be reduced to the HSP in the dihedral group (assuming it is solved by the standard method discussed below) [44]. Unfortunately, no efficient algorithms are known for these two instances of the nonabelian hidden subgroup problem. However, some partial progress has been made: there is a subexponential time algorithm for the dihedral HSP [37, 45], and it is known how to solve the HSP efficiently for a variety of other nonabelian groups [2, 20, 22, 24, 30, 33, 40].

In the hidden subgroup problem for a group  $G$ , we are given black-box access to a function  $f : G \rightarrow S$ , where  $S$  is some finite set. We say that  $f$  hides a subgroup  $H \leq G$  provided  $f(g) = f(g')$  if and only if  $g^{-1}g' \in H$  (or in other words, provided  $f$  is constant on left cosets

---

\*amchilds@caltech.edu

†a.harrow@bris.ac.uk

‡wocjan@cs.ucf.edu

of  $H$  in  $G$  and distinct on different left cosets). The goal is to determine  $H$  (say, in terms of a generating set) as quickly as possible. In particular, we say that an algorithm for the HSP in  $G$  is efficient if it runs in time  $\text{poly}(\log |G|)$ .

Nearly all quantum algorithms for the hidden subgroup problem make use of the so-called *standard method*, in which one queries the function  $f$  on a uniform superposition of group elements and then discards the function value, giving a *coset state*

$$|gH\rangle := \frac{1}{\sqrt{|H|}} \sum_{h \in H} |gh\rangle \quad (1)$$

for some unknown, uniformly random  $g \in G$ . In other words, the state is described by the density matrix

$$\rho_H := \frac{1}{|G|} \sum_{g \in G} |gH\rangle \langle gH| \quad (2)$$

$$= \frac{1}{|G|} \sum_{h \in H} R(h) \quad (3)$$

(sometimes called a *hidden subgroup state*), where  $R$  is the *right regular representation* of  $G$ , satisfying

$$R(g)|g'\rangle = |g'g^{-1}\rangle \quad (4)$$

for all  $g, g' \in G$ . Now the hidden subgroup problem is reduced to the problem of performing a measurement to distinguish the states  $\rho_H$  for the various possible  $H \leq G$ .

The symmetry of the hidden subgroup state can be exploited using Fourier analysis. In particular, the group algebra  $\mathbb{C}G$  decomposes under the left and right multiplication actions of  $G$  (which commute, and indeed generate each other's commutants) as

$$\mathbb{C}G \cong \bigoplus_{\sigma \in \hat{G}}^{G \times G} \mathcal{V}_\sigma \otimes \mathcal{V}_\sigma^* \quad (5)$$

where  $\hat{G}$  denotes a complete set of irreducible representations (or *irreps*) of  $G$ , and  $\mathcal{V}_\sigma$  and  $\mathcal{V}_\sigma^*$  are the (row and column, respectively) subspaces acted on by  $\sigma \in \hat{G}$ . The unitary transformation that relates the standard basis for  $\mathbb{C}G$  and the basis for the spaces  $\mathcal{V}_\sigma \otimes \mathcal{V}_\sigma^*$  is the Fourier transform, which can be carried out efficiently for most groups of interest [9, 15, 28, 39].

Since the hidden subgroup state  $\rho_H$  is invariant under the left multiplication action of  $G$ , the decomposition (5) shows that it is block diagonal in the Fourier basis, with blocks labeled by the irreps  $\sigma \in \hat{G}$ . For each  $\sigma$ , there is a  $\dim \mathcal{V}_\sigma \times \dim \mathcal{V}_\sigma$  block that appears  $\dim \mathcal{V}_\sigma$  times (or in other words, the state is maximally mixed in the row space). Thus, without loss of information (or in Kuperberg's words, "sacrificing no entropy"), we can measure the irrep name  $\sigma$  and discard the information about which  $\sigma$ -isotypic block occurred (sometimes referred to as "discarding the row label").

The process of measuring the irrep name  $\sigma$  is referred to as *weak Fourier sampling*. If  $G$  is abelian, then all its irreps are 1-dimensional. In this case, weak Fourier sampling is a complete measurement, and all that remains is to use the statistics of the irrep name to infer the hidden subgroup. However, for the nonabelian HSP, some (indeed, typically most) irreps are more than 1-dimensional, so weak Fourier sampling is an incomplete measurement. Indeed, for most nonabelian groups, it turns out that weak Fourier sampling alone produces insufficient information to identify the hidden subgroup. In particular, this is true of the HSP relevant to graph isomorphism [24, 30] as well as the dihedral HSP. To obtain further information about the hidden subgroup, one must perform a refined measurement inside the  $(\dim \mathcal{V}_\sigma)$ -dimensional subspace that remains after weak Fourier sampling gives the result  $\sigma$ . The process of making such a refined measurement is referred to as *strong Fourier sampling*. There are many possible ways to measure inside  $\mathcal{V}_\sigma$ , especially when it is high-dimensional, and much work has been done on the possible measurements for particular nonabelian groups.

Of course, with either weak or strong Fourier sampling, a single hidden subgroup state is not sufficient to determine  $H$ : one must repeat the sampling procedure to obtain statistics. However, if one simply repeats strong Fourier sampling a polynomial number of times, there are some groups—in particular, the symmetric group—in which hidden subgroups cannot be found, even if the measurement is allowed to be chosen adaptively [41]. Thus, to solve the HSP in general, one must perform a joint measurement on  $k = \text{poly}(\log |G|)$  copies of the hidden subgroup state,  $\rho_H^{\otimes k}$ . In fact, there are groups—again including the symmetric group—for which the measurement must be entangled across  $\Omega(\log |G|)$  copies [29]. Thus the difficulty of the general HSP may be attributed at least in part to that fact that highly entangled measurements are required (although it should be noted that there are some groups, such as the dihedral group, where single-register measurements are at least information-theoretically sufficient [18], yet an efficient algorithm remains elusive). While it is known that  $O(\log |G|)$  copies are always information-theoretically sufficient [19] (so that, in particular, the query complexity of the HSP is polynomial), there are many groups for which it is not known how to efficiently extract the identity of the hidden subgroup.

Although previous work on the hidden subgroup problem has focused almost exclusively on Fourier sampling, it turns out that there is another kind of measurement that can also be performed without loss of information. The idea is to exploit the fact that the state  $\rho_H^{\otimes k}$  is symmetric under permutations of the  $k$  registers. Thus, we should consider the decomposition of the space  $(\mathbb{C}G)^{\otimes k}$  afforded by *Schur duality* [23], which decomposes  $k$  copies of a  $d$ -dimensional space as

$$(\mathbb{C}^d)^{\otimes k} \xrightarrow{\mathcal{S}_k \times \mathcal{U}_d} \bigoplus_{\lambda \vdash k} \mathcal{P}_\lambda \otimes \mathcal{Q}_\lambda^d \quad (6)$$

where the symmetric group  $\mathcal{S}_k$  acts to permute the  $k$  registers and the unitary group  $\mathcal{U}_d$  acts identically on each register. Here the subspaces  $\mathcal{P}_\lambda$  and  $\mathcal{Q}_\lambda^d$  correspond to irreps of  $\mathcal{S}_k$  and  $\mathcal{U}_d$ , respectively. They are labeled by partitions  $\lambda$  of  $k$  (denoted  $\lambda \vdash k$ ), i.e.,  $\lambda = (\lambda_1, \lambda_2, \dots)$  where  $\lambda_1 \geq \lambda_2 \geq \dots$  and  $\sum_j \lambda_j = k$ . Note that we can restrict our attention to partitions with at most  $d$  parts, since  $\dim \mathcal{Q}_\lambda^d = 0$  if  $\lambda_{d+1} > 0$ .

Since  $\rho_H^{\otimes k}$  is invariant under the action of  $\mathcal{S}_k$ , the decomposition (6) shows that it is block diagonal in the Schur basis with blocks labeled by  $\lambda \vdash k$ . For each  $\lambda$ , there is a  $\dim \mathcal{Q}_\lambda^{|G|} \times \dim \mathcal{Q}_\lambda^{|G|}$  block that appears  $\dim \mathcal{P}_\lambda$  times (or in other words, the state is maximally mixed in the permutation space). Thus, no information is lost if we measure the partition  $\lambda$  and discard the permutation register. By analogy to weak Fourier sampling, we refer to the process of measuring  $\lambda$  as *weak Schur sampling*. This is a natural measurement to consider not only because it can be performed without loss of information, but also because it is a joint measurement of all  $k$  registers, and we know that some measurement of this kind is required to solve the general HSP. Unfortunately, we will see in Section 2 (and see also Corollary 4 below) that weak Schur sampling with  $k = \text{poly}(\log |G|)$  provides insufficient information to solve the HSP unless the hidden subgroup is very large (in which case the problem is easy, even for a classical computer).

In fact, since both weak Fourier sampling and weak Schur sampling can be performed without loss of information, it is possible to perform both measurements simultaneously (with the caveat that one must discard the irrelevant information about the order in which the irreps of  $G$  appear). Even though the statistics of the irrep name  $\sigma$  and the partition  $\lambda$  do not provide enough information to identify the hidden subgroup, this does not preclude the possibility that their joint distribution is more informative. However, we will see in Section 3 that unless we are likely to see the same representation more than once under weak Fourier sampling (which is typically not the case), the Fourier and Schur distributions are nearly uncorrelated. Formally, we have

**Theorem 1 (Failure of weak Fourier-Schur sampling).** *The probability that weak Fourier-Schur sampling (defined in Section 3) applied to  $\rho_H^{\otimes k}$  (defined in (3)) provides a result that depends on  $|H|$  is at most  $k^2 |H| d_{\max}^2 / |G|$ , where  $d_{\max}$  is the largest dimension of an irrep of  $G$ .*

This implies that  $k$  needs to be large for most cases of interest, including the dihedral and

symmetric groups.

**Corollary 2 (Weak Fourier-Schur sampling on  $\mathcal{D}_N$  and  $\mathcal{S}_n$ ).** (a) *Weak Fourier-Schur sampling on the dihedral group  $\mathcal{D}_N$  cannot distinguish the trivial subgroup from a hidden reflection with constant advantage unless  $k = \Omega(\sqrt{N})$ .* (b) *Weak Fourier-Schur sampling on the symmetric group  $\mathcal{S}_n$  or on the wreath product  $\mathcal{S}_n \wr \mathbb{Z}_2$  cannot distinguish the trivial subgroup from an order 2 subgroup with constant advantage unless  $k = \exp(\Omega(\sqrt{n}))$ .*

While it is unfortunate that weak Fourier-Schur sampling typically cannot provide enough information to find a hidden subgroup, it remains possible that Schur duality could be a helpful tool for solving the HSP. Just as weak Fourier sampling breaks the space into smaller subspaces in which one can perform strong Fourier sampling (or a multi-register generalization thereof), weak Fourier-Schur sampling provides an even finer decomposition into subspaces. Since the latter decomposition produces subspaces that are entangled across all  $k$  registers, it may be especially useful for HSPs where we know that entangled measurements are necessary (or for which we would like to implement optimal measurements, which are typically entangled [2, 3]).

The proof that weak Schur sampling fails is based on the simple observation that distinguishing the trivial subgroup from a subgroup of order  $|H|$  in this way requires one to solve the problem of distinguishing 1-to-1 from  $|H|$ -to-1 functions on  $G$ , i.e., the  $|H|$ -collision problem for a list of size  $|G|$ . Since there is an  $\Omega(\sqrt[3]{|G|/|H|})$  quantum lower bound for this problem [1],  $\text{poly}(\log |G|)$  registers are insufficient. In fact, the problem resulting from the HSP is potentially harder, since the basis in which the collisions occur is inaccessible to the Schur measurement. This naturally leads to the notion of a *quantum collision problem*, and raises the question of how quickly it can be solved on a quantum computer. In Section 4, we define a sampling version of the quantum  $r$ -collision problem and use results on the asymptotics of the Plancherel measure on the symmetric group to prove that  $k = \Theta(d/r)$  registers are necessary and sufficient to solve it. In particular, we prove

**Theorem 3 (Quantum collision sampling problem).** *Given  $\rho^{\otimes k}$ , distinguishing between [case A]  $\rho = I/d$  and [case B]  $\rho^2 = \rho/\frac{d}{r}$  (i.e.,  $\rho$  is proportional to a projector of rank  $d/r$ ) is possible with success probability  $1 - \exp(-\Theta(kr/d))/2$ . In particular, success probability  $\frac{1}{2} + \Omega(1)$  is possible if and only if  $k = \Theta(d/r)$ .*

In addition to providing the first results on estimation of the spectrum of a quantum state in the regime where  $k \ll d^2$ , this gives tight estimates of the effectiveness of weak Schur sampling, which we see requires an exponentially large (in  $\log |G|$ ) number of copies to be successful.

**Corollary 4 (Failure of Weak Schur sampling).** *Applying Weak Schur sampling to  $\rho_H^{\otimes k}$  (where  $\rho_H$  is defined in (3)) requires  $k = \Theta(|G|/r)$  to distinguish the case  $|H| \geq r$  from the case  $H = \{1\}$  with constant advantage.*

The connection between Theorem 3 and Corollary 4 is explained in Section 2, while Theorem 3 is proved in Section 4.

Finally, in Section 4.2 we introduce a black box version of the quantum collision problem. We show that it can be solved using  $O(\sqrt[3]{d/r} \log d/r)$  queries, nearly matching the query lower bound from the classical problem (although the best algorithm we are aware of has a running time of  $O(\sqrt{d/r} \log d/r)$ ). We also discuss some alternative formulations of the quantum collision problem that are less closely connected to the HSP, but that may be of independent interest.

## 2 Weak Schur sampling

We begin by considering only the permutation symmetry of  $k$  independent copies of the  $d$ -dimensional hidden subgroup states, without taking into account symmetry resulting from the group  $G$ . In other words, we will consider only the Schur decomposition (6), and we will perform *weak Schur sampling*, i.e., a measurement of the partition  $\lambda$ .

The projector onto the subspace labeled by a particular  $\lambda \vdash k$  is

$$\Pi_\lambda := \frac{\dim \mathcal{P}_\lambda}{k!} \sum_{\pi \in \mathcal{S}_k} \chi_\lambda(\pi) P(\pi) \quad (7)$$

(see for example [49, Theorem 8]), where  $\chi_\lambda$  is the character of the irrep of  $\mathcal{S}_k$  labeled by  $\lambda$  and  $P$  is the (reducible) representation of  $\mathcal{S}_k$  that acts to permute the  $k$  registers, i.e.,

$$P(\pi)|i_1\rangle|i_2\rangle\cdots|i_k\rangle = |i_{\pi^{-1}(1)}\rangle|i_{\pi^{-1}(2)}\rangle\cdots|i_{\pi^{-1}(k)}\rangle \quad (8)$$

for all  $i_1, \dots, i_k \in \{1, \dots, d\}$ . For any  $d^k$ -dimensional density matrix  $\gamma$ , the probability distribution under weak Schur sampling is

$$\Pr(\lambda|\gamma) = \text{tr}(\Pi_\lambda \gamma). \quad (9)$$

To use weak Schur sampling as part of a quantum algorithm, it is important that the measurement of  $\lambda$  can be carried out efficiently. The simplest implementation of the complete Schur transform [5], which fully resolves the subspaces  $\mathcal{P}_\lambda$  and  $\mathcal{Q}_\lambda^d$ , runs in time  $\text{poly}(k, d)$ , and thus is inefficient when  $d$  is exponentially large, as in the hidden subgroup problem. It can be modified to run in time  $\text{poly}(k, \log d)$  either by a relabeling trick [31, footnote in Section 8.1.2] or by using so-called *generalized phase estimation* [4, 31] (which may also be viewed as a generalization of the well-known swap test [8, 13]). Generalized phase estimation only allows us to measure  $\lambda$ , but for weak Schur sampling this is all we need. In this procedure, we prepare an ancilla register in the state  $\frac{1}{\sqrt{k!}} \sum_{\pi \in \mathcal{S}_k} |\pi\rangle$ , use it to perform a conditional permutation  $P(\pi)$  on the input state  $\gamma$ , and then perform an inverse Fourier transform over  $\mathcal{S}_k$  [9] on the ancilla register. Measurement of the ancilla register will then yield  $\lambda \in \hat{\mathcal{S}}_k$ , interpreted as a partition of  $k$ , distributed according to (9).

The distribution of  $\lambda$  according to weak Schur sampling is invariant under the actions of the permutation and unitary groups, since these groups act only within the subspaces  $\mathcal{P}_\lambda$  and  $\mathcal{Q}_\lambda^d$ , respectively. In other words, for any  $d^k$ -dimensional density matrix  $\gamma$ , we have

$$\Pr(\lambda|U^{\otimes k} \gamma U^{\dagger \otimes k}) = \Pr(\lambda|\gamma) \quad \forall U \in \mathcal{U}_d \quad (10)$$

$$\Pr(\lambda|P(\pi) \gamma P(\pi^{-1})) = \Pr(\lambda|\gamma) \quad \forall \pi \in \mathcal{S}_k. \quad (11)$$

(It is straightforward to verify these invariances by directly computing  $[\Pi_\lambda, U^{\otimes k}] = [\Pi_\lambda, P(\pi)] = 0$  and using the cyclic property of the trace.) In particular, (10) implies that for  $\gamma = \rho^{\otimes k}$ , the distribution according to weak Schur sampling depends only on the spectrum of  $\rho$ . (In fact, it turns out that simply normalizing  $\lambda$  provides a good estimate of the spectrum of  $\rho$ , provided  $k$  is large [36].)

Now it is easy to see that weak Schur sampling on  $k = \text{poly}(\log |G|)$  copies of the hidden subgroup state  $\rho_H$  provides insufficient information to solve the hidden subgroup problem. The hidden subgroup state  $\rho_H$  is proportional to a projector of rank  $|G|/|H|$ , as can be seen by computing

$$\rho_H^2 = \frac{1}{|G|^2} \sum_{h, h' \in H} R(hh') = \frac{|H|}{|G|} \rho_H. \quad (12)$$

Since the distribution of measurement outcomes  $\Pr(\lambda|\rho_H^{\otimes k})$  depends only on the spectrum of  $\rho_H$ , and the spectrum of  $\rho_H$  depends only on the order of the hidden subgroup, it is clear that different subgroups of the same order cannot be distinguished by weak Schur sampling. Nevertheless, we might hope to distinguish the trivial hidden subgroup from a hidden subgroup of order  $|H| \geq 2$ , which would be sufficient to solve the HSP relevant to graph isomorphism, for example. However, even this task requires an exponential number of hidden subgroup states, which can be seen as follows.

Suppose that weak Schur sampling could distinguish between hidden subgroup states corresponding to  $H = \{1\}$  and some particular  $H$  of order  $|H| \geq 2$ . Since the distribution of  $\lambda$

depends only on the spectrum of the state, this would mean that we could distinguish  $k$  copies of the maximally mixed state  $I_{|G|}/|G|$ , where  $I_d$  is the  $d \times d$  identity matrix, from  $k$  copies of the state  $J_{|G|/|H|}/(|G|/|H|)$ , where  $J_{d'}$  is a projector onto an arbitrary subspace of dimension  $d'$ . This in turn would imply that we could distinguish 1-to-1 functions from  $|H|$ -to-1 functions using  $k$  queries of the function. Then the quantum lower bound for the  $|H|$ -collision problem [1] shows that  $k = \Omega(\sqrt[3]{|G|/|H|})$  copies are required.

Of course, this does not mean that  $O(\sqrt[3]{|G|/|H|})$  copies are sufficient. In fact, it turns out that a linear number of copies is both necessary and sufficient, as we will show by a more careful analysis in Section 4. There we will prove Theorem 3, which by the arguments of this section implies Corollary 4.

### 3 Weak Fourier-Schur sampling

In the previous section, we showed that weak Schur sampling provides insufficient information to efficiently solve the hidden subgroup problem. However, even though weak Fourier sampling typically also does not provide enough information, it is conceivable that the joint distribution of the two measurements could be substantially more informative. In this section, we will see that this is not the case: provided weak Fourier sampling fails, so does weak Fourier-Schur sampling.

Because neither measurement constitutes a loss of information, it is in principle possible to perform both weak Fourier sampling and weak Schur sampling simultaneously. If we perform weak Fourier sampling in the usual way, measuring the irrep label for each register, then we will typically obtain a state that is no longer permutation invariant. However, since the irrep labels are identically distributed for each register, the order in which the irreps appear carries no information. Only the *type* of the irreps, i.e., the number of times each irrep appears, is important. Thus, it is sufficient to perform what we might call *weak Fourier type sampling*, in which we only measure the irrep type. Equivalently, we could perform complete weak Fourier sampling and then randomly permute the  $k$  registers, thereby discarding the (irrelevant) information about the order of the irreps, and producing a permutation-invariant state to use for weak Schur sampling. (In fact, since the distribution under weak Schur sampling has the permutation invariance (11), it is not actually necessary to explicitly symmetrize the state.)

We begin by performing weak Fourier sampling. The hidden subgroup state  $\rho_H$  defined in (3) has the following block structure in the Fourier basis:

$$\rho_H \cong \frac{1}{|G|} \bigoplus_{\sigma \in \hat{G}} I_{\dim \mathcal{V}_\sigma} \otimes \sum_{h \in H} \sigma(h)^* \quad (13)$$

$$=: \sum_{\sigma \in \hat{G}} \Pr(\sigma) \frac{I_{\dim \mathcal{V}_\sigma}}{\dim \mathcal{V}_\sigma} \otimes \rho_{H,\sigma}. \quad (14)$$

Here the probability of observing the irrep  $\sigma$  is

$$\Pr(\sigma) = \frac{\dim \mathcal{V}_\sigma}{|G|} \sum_{h \in H} \chi_\sigma(h)^*, \quad (15)$$

and the state conditioned on this observation is  $\rho_{H,\sigma}$ . Weak Fourier sampling consists of measuring the label of the irrep. Assume we have observed  $\sigma$ . Then the resulting density operator is

$$\rho_{H,\sigma} = \frac{1}{\sum_{h \in H} \chi_\sigma(h)} \sum_{h \in H} |\sigma\rangle\langle\sigma| \otimes \sigma(h)^*. \quad (16)$$

We explicitly include the label  $|\sigma\rangle$  to reflect the fact that  $\sigma$  was observed.

Repeating weak Fourier sampling  $k$  times, we obtain the state

$$\rho_{H,\underline{\sigma}} = \rho_{H,\sigma_1} \otimes \cdots \otimes \rho_{H,\sigma_k}, \quad (17)$$

where  $\underline{\sigma} := (\sigma_1, \sigma_2, \dots, \sigma_k) \in \hat{G}^k$  may be viewed either as the actual outcome of the  $k$  instances of weak Fourier sampling, or merely as a representative of the irrep type, as discussed above. Given this state, the conditional probability of observing the partition  $\lambda$  is

$$\Pr(\lambda|\underline{\sigma}) = \text{tr}(\Pi_\lambda \rho_{H,\underline{\sigma}}) \quad (18)$$

$$= \frac{\dim \mathcal{P}_\lambda}{k!} \sum_{\pi \in \mathcal{S}_k} \chi_\lambda(\pi) \text{tr}[P(\pi) \rho_{H,\underline{\sigma}}]. \quad (19)$$

Note that  $\text{tr}[P(\pi) \rho_{H,\underline{\sigma}}] = 0$  if  $\pi(\underline{\sigma}) \neq \underline{\sigma}$ , where  $\pi(\underline{\sigma}) = (\sigma_{\pi^{-1}(1)}, \dots, \sigma_{\pi^{-1}(k)})$ .

Now we are ready to prove Theorem 1:

*Proof of Theorem 1.* Let us assume that  $\underline{\sigma}$  is multiplicity-free, i.e., that all the  $\sigma_i$ 's are different. In this case the traces are zero for all  $\pi \neq 1$ , where 1 is the identity element of  $\mathcal{S}_k$ . Then

$$\Pr(\lambda|\underline{\sigma}) = \frac{\dim \mathcal{P}_\lambda}{k!} \chi_\lambda(1) \text{tr} \rho_{H,\underline{\sigma}} = \frac{(\dim \mathcal{P}_\lambda)^2}{k!}, \quad (20)$$

which is nothing but the Plancherel distribution over  $\hat{\mathcal{S}}_k$ , and which in particular is independent of the hidden subgroup  $H$ . This shows that we cannot extract any information about  $H$  provided that we have obtained a multiplicity-free  $\underline{\sigma}$ .

It remains to show that the probability of obtaining a type  $\underline{\sigma}$  that is not multiplicity-free is vanishingly small. This can be seen as follows. From (15), we have

$$\Pr(\sigma) \leq \frac{(\dim \mathcal{V}_\sigma)^2}{|G|} |H| \quad (21)$$

since the absolute value of the character  $\chi_\sigma$  at any group element can be at most  $\dim \mathcal{V}_\sigma$ . If we perform weak Fourier sampling on two copies of  $\rho_H$ , then the probability of obtaining the same irrep twice is

$$\Pr(2 \text{ irreps same}) = \sum_{\sigma \in \hat{G}} \Pr(\sigma)^2 \leq \max_{\sigma \in \hat{G}} \Pr(\sigma) \leq \frac{d_{\max}^2}{|G|} |H|, \quad (22)$$

where  $d_{\max} := \max_{\sigma \in \hat{G}} \dim \mathcal{V}_\sigma$  denotes the maximal dimension of any irrep of  $G$ . For  $k$  copies, we can upper bound the probability that any two irreps out of the  $k$  observed irreps are the same by a union bound. We find

$$\Pr(\text{any 2 irreps same}) \leq \binom{k}{2} \Pr(2 \text{ irreps same}) \leq \binom{k}{2} \frac{d_{\max}^2}{|G|} |H|. \quad (23)$$

Therefore, provided  $|H|$  and  $d_{\max}$  are not too large, the probability of obtaining the same irrep twice is small. This concludes the proof of Theorem 1.  $\square$

We now focus on the important special cases of the dihedral and symmetric groups to prove Corollary 2.

*Proof of Corollary 2.* When  $G$  is the dihedral group  $\mathcal{D}_N$  of order  $2N$ , we have  $d_{\max} = 2$ , so (23) shows that weak Fourier-Schur sampling cannot distinguish the trivial subgroup from a hidden reflection (a subgroup of order 2) with constant advantage unless  $k = \Omega(\sqrt{N})$ . Similarly, for the symmetric group  $G = \mathcal{S}_n$ , we have  $d_{\max} \leq \sqrt{n!} \exp(-c\sqrt{n})$  for some constant  $c$  [53], which implies that weak Fourier-Schur sampling cannot solve the HSP relevant to graph isomorphism unless  $k = \exp(\Omega(\sqrt{n}))$ . Finally, we can express the irreps of the wreath product  $\mathcal{S}_n \wr \mathbb{Z}_2$  in terms of the irreps of  $\mathcal{S}_n$  [49, Section 8.2], implying that  $d_{\max} \leq 2n! \exp(-2c\sqrt{n})$ . Since  $|\mathcal{S}_n \wr \mathbb{Z}_2| = 2(n!)^2$ , we again have exponentially small distinguishability unless  $k = \exp(\Omega(\sqrt{n}))$ .  $\square$

In [14] two of us considered an alternative quantum approach to the graph isomorphism problem based on the nonabelian hidden shift problem. It can be shown that weak Fourier-Schur sampling fails for similar reasons when applied to hidden shift states instead of hidden subgroup states. The hidden shift states have a similar block structure, where the blocks correspond to irreps of  $G$  (although in the hidden shift case, the distribution of  $\sigma$  is exactly Plancherel over  $\hat{G}$ , so weak Fourier sampling trivially fails). Again, the probability of seeing the same irrep twice is exponentially small, and unless some irrep appears twice, the conditional distribution of  $\lambda$  is equal to the Plancherel distribution over  $\hat{S}_k$ .

## 4 The quantum collision problem

In Section 2, we saw that weak Schur sampling cannot efficiently solve the hidden subgroup problem since such a solution would imply an efficient solution of the collision problem. In fact, the problem faced by weak Schur sampling is considerably more difficult, since no information is available about the basis in which the collisions occur. This motivates quantum generalizations of the usual (i.e., classical) collision problem, which we study in this section.

The classical  $r$ -collision problem is the problem of determining whether a black box function with  $d$  inputs (such that  $r$  divides  $d$ ) is either 1-to-1 or  $r$ -to-1. This problem has classical (randomized) query complexity  $\Theta(\sqrt{d/r})$ —as evidenced by the well-known birthday problem—and quantum query complexity  $\Theta(\sqrt[3]{d/r})$  [1, 11]. The classical algorithm is quite simple: after querying the function on  $O(\sqrt{d/r})$  randomly chosen inputs, there is a reasonable probability that a collision will appear, provided one exists. The quantum algorithm is slightly more subtle, making use of Grover’s algorithm for unstructured search [25]. In particular, while the classical algorithm queries the black box non-adaptively, it is essential for the quantum algorithm to make adaptive queries.

We begin in Section 4.1 by considering what one might call the *quantum  $r$ -collision sampling problem*, which is simply the problem of deciding whether one has  $k$  copies of the  $d$ -dimensional maximally mixed state or of a state that is maximally mixed on an unknown subspace of dimension  $d/r$ . This is exactly the problem faced by the weak Schur sampling approach to the hidden subgroup problem, so our results on the quantum collision sampling problem allow us to put tight bounds on the effectiveness of weak Schur sampling for the HSP. It turns out that  $k = \Theta(d/r)$  copies are necessary and sufficient to distinguish these two cases with constant advantage. Interestingly, this measurement is entangled across all  $k$  registers, and we do not know whether its performance can be matched by an adaptive sequence of single-register measurements, as is the case for asymptotically large  $k$  [6]. However, we show that non-adaptively measuring  $k$  sets of  $O(1)$  registers each is much less effective, succeeding only if  $k = \Theta(d^2/r^2)$ .

A complete definition of the quantum collision problem requires us to specify a unitary black box that hides the function, and that allows us to make adaptive queries. We propose one such definition in Section 4.2, and show that the resulting quantum  $r$ -collision problem can be solved in  $O(\sqrt[3]{d/r} \log d/r)$  queries, nearly matching the  $\Omega(\sqrt[3]{d/r})$  lower bound from the classical collision problem. We also mention some possible alternative formulations of the problem and comment on their complexities.

### 4.1 The quantum collision sampling problem

This section is devoted to the proof of Theorem 3. We consider the sampling version of the quantum collision problem, in which we are given  $k$  samples of a quantum state that is either maximally mixed on a Hilbert space of dimension  $d$ , or is maximally mixed on an unknown (but fixed) subspace of dimension  $d/r$ . We begin by showing the equivalence of this problem with some related problems.



**Lemma 5.** *The following tasks are equivalent:*

(a) *Distinguishing [case A]  $(I/d)^{\otimes k}$  from [case B]*

$$\int dU \left( \sum_{j=1}^{d/r} U|j\rangle\langle j|U^\dagger \right)^{\otimes k} \quad (24)$$

where  $dU$  is the Haar measure on the  $d$ -dimensional unitary group  $\mathcal{U}_d$ .

- (b) *Distinguishing [case A]  $(I/d)^{\otimes k}$  from [case B]  $\rho^{\otimes k}$  for arbitrary  $\rho$  satisfying  $\rho^2 = \rho/d$  (i.e., proportional to a projector of rank  $d/r$ ).*
- (c) *Distinguishing [case A]  $\rho_1^{\otimes k}$  from [case B]  $\rho_2^{\otimes k}$ , where  $\rho_1, \rho_2$  are density matrices on  $\mathbb{C}^{d'}$ ,  $d' \geq d$ ,  $\rho_1^2 = \rho_1/d$  and  $\rho_2^2 = \rho_2/d$ .*

Moreover, the optimal solution to each problem is to perform weak Schur sampling (i.e., the measurement with operators  $\{\Pi_\lambda\}_{\lambda \vdash k}$  defined in (7)) and then to perform classical processing on the measurement outcome. Both the quantum measurement and the classical processing can be performed in time  $\text{poly}(k, \log d)$ .

*Proof.* First note that (a) reduces to (b), which in turn reduces to (c). Thus equivalence of all three will follow from reducing (c) to (a). To do this, we will demonstrate an optimal strategy for (a) and show that it performs equally well when applied to (c).

Both states in (a) are invariant under the actions of  $\mathcal{S}_k$  and  $\mathcal{U}_d$ , so they can each be written in the form  $\sum_\lambda \Pr(\lambda|\gamma) \Pi_\lambda / \text{tr} \Pi_\lambda$ , where  $\gamma$  is the state appearing either in case A or case B, the  $\Pi_\lambda$  are defined according to (7), and  $\Pr(\lambda|\gamma) = \text{tr} \Pi_\lambda \gamma$ . Thus weak Schur sampling (followed by classical processing of the outcome  $\lambda$ ) is an optimal measurement strategy for problem (a).

Now observe from (7) that  $\Pi_\lambda$  can be defined as a weighted sum of permutations, and thus  $\text{tr} \Pi_\lambda \rho^{\otimes k}$  depends on  $\rho$  only via the quantities  $\text{tr} \rho, \text{tr} \rho^2, \dots, \text{tr} \rho^k$  that appear when evaluating  $\text{tr} P(\pi) \rho^{\otimes k}$  for  $\pi \in \mathcal{S}_k$ . This means that  $\Pr(\lambda|\gamma)$  is unchanged if we either embed  $(\mathbb{C}^d)^{\otimes k}$  into  $(\mathbb{C}^{d'})^{\otimes k}$  for  $d' \geq d$ , or if we replace  $\rho^{\otimes k}$  with  $(U\rho U^\dagger)^{\otimes k}$ , or indeed with  $\mathbf{E}_U(U\rho U^\dagger)^{\otimes k}$  for any distribution over unitaries  $U$ . This implies that  $\Pr(\lambda|\gamma)$  is the same for tasks (a), (b) and (c), and that any solution for (a) implies an equally effective solution for (c).

Finally, the measurement  $\{\Pi_\lambda\}_{\lambda \vdash k}$  (i.e., weak Schur sampling) can be efficiently implemented using generalized phase estimation, as described in Section 2. The classical processing consists of comparing the conditional probabilities  $\text{tr} \Pi_\lambda \rho^{\otimes k}$ , for which we give efficiently calculable formulas below.  $\square$

Now we turn to the calculation of  $\Pr(\lambda) = \text{tr} \Pi_\lambda \rho^{\otimes k}$  for  $\rho$  satisfying  $\rho^2 = \rho/d$  (and embedded into an ambient space of arbitrary dimension  $d' \geq d$ ). We call the resulting distribution of  $\lambda \vdash k$  the *Schur distribution*,  $\text{Schur}(k, d)$ , in which  $\lambda$  appears with probability

$$\Pr(\lambda) = \frac{\dim \mathcal{P}_\lambda \dim \mathcal{Q}_\lambda^d}{d^k} \quad (25)$$

$$= \frac{(\dim \mathcal{P}_\lambda)^2}{k!} \prod_{(i,j) \in \lambda} \left( 1 + \frac{j-i}{d} \right), \quad (26)$$

where in the second line we have used Stanley's formula for  $\dim \mathcal{Q}_\lambda^d$  [51], interpreting  $\lambda$  as a Young diagram, where  $(i, j) \in \lambda$  iff  $1 \leq j \leq \lambda_i$ .

Our goal is to show that the distributions  $\text{Schur}(k, d)$  and  $\text{Schur}(k, d/r)$  are close unless  $k$  is  $\Omega(d/r)$ . We will do this by comparing the Schur distribution to the Plancherel distribution over  $\lambda \in \hat{\mathcal{S}}_k$ ,  $\text{Planch}(k)$ , for which

$$\Pr(\lambda) = \frac{(\dim \mathcal{P}_\lambda)^2}{k!}. \quad (27)$$

Note that we can think of  $\text{Planch}(k)$  as the  $d \rightarrow \infty$  limit of  $\text{Schur}(k, d)$ .

The key technical part of the proof of Theorem 3 is the following lemma:

**Lemma 6.** *If  $2 \leq k \leq d$  then*

$$\frac{k}{36d} \leq \|\text{Schur}(k, d) - \text{Planch}(k)\|_1 \leq \sqrt{2} \frac{k}{d}. \quad (28)$$

We note in passing that the upper bound can be amplified when  $k \geq d$  as follows:

**Corollary 7.**

$$\frac{1}{2} \|\text{Schur}(k, d) - \text{Planch}(k)\|_1 \geq 1 - e^{-\frac{1}{10368}(\frac{k}{d}-1)}. \quad (29)$$

*Proof of Corollary 7.* For density matrices  $\rho, \sigma$ , we define the fidelity  $F(\rho, \sigma) := \|\sqrt{\rho}\sqrt{\sigma}\|_1^2$  and recall the relations [21]

$$1 - \sqrt{F(\rho, \sigma)} \leq \frac{1}{2} \|\rho - \sigma\|_1 \leq \sqrt{1 - F(\rho, \sigma)}, \quad (30)$$

where the second inequality is saturated when  $\rho$  and  $\sigma$  are pure states. Now use Lemma 6 to obtain

$$\sqrt{1 - F(\text{Schur}(d, d), \text{Planch}(d))} \geq \frac{1}{2} \|\text{Schur}(d, d) - \text{Planch}(d)\|_1 \geq \frac{1}{72}. \quad (31)$$

Next,

$$F(\text{Schur}(k, d), \text{Planch}(k)) \leq F(\text{Schur}(d, d), \text{Planch}(d))^{\lfloor \frac{k}{d} \rfloor} \quad (32)$$

$$\leq \left(1 - \frac{1}{5184}\right)^{\lfloor \frac{k}{d} \rfloor} \leq e^{-\frac{1}{5184}(\frac{k}{d}-1)}, \quad (33)$$

using first that fidelity is nondecreasing under quantum operations and is multiplicative under tensor products, then (31), and then the inequality  $1 - x \leq e^{-x}$ . Finally,

$$\frac{1}{2} \|\text{Schur}(k, d) - \text{Planch}(k)\|_1 \geq 1 - \sqrt{F(\text{Schur}(k, d), \text{Planch}(k))} \geq 1 - e^{-\frac{1}{10368}(\frac{k}{d}-1)}, \quad (34)$$

which completes the proof.  $\square$

To derive Theorem 3 from Lemma 6, we need one additional fact.

**Lemma 8 (Monotonicity).** *For any integers  $k, d_1, d_2 \geq 1$  and  $r \geq 2$ ,*

$$\|\text{Schur}(k, d_1) - \text{Schur}(k, d_2)\|_1 \geq \|\text{Schur}(k, rd_1) - \text{Schur}(k, rd_2)\|_1. \quad (35)$$

*Proof.* Define  $\sigma_d$  (for  $d$  one of  $d_1, d_2, rd_1, rd_2$ ) as the state

$$\sigma_d := \int dU \left( \sum_{j=1}^d U|j\rangle\langle j|U^\dagger \right)^{\otimes k} \quad (36)$$

where  $dU$  is the Haar measure on the  $d'$ -dimensional unitary group  $\mathcal{U}_{d'}$  and  $d'$  is some arbitrary integer satisfying  $d' \geq r \cdot \max(d_1, d_2)$ . Let  $\mathcal{N}$  be the quantum operation which takes as input a state on  $(\mathbb{C}^{d'})^{\otimes k}$ , adjoins the maximally mixed state  $(I_r/r)^{\otimes k}$ , draws a uniform  $U \in \mathcal{U}_{rd'}$ , applies  $U^{\otimes k}$ , and finally outputs a state on  $(\mathbb{C}^{rd'})^{\otimes k}$ . Adjoining  $I_r/r$  will map a state proportional to a rank  $d$  projector to a state proportional to a rank  $rd$  projector. Thus,  $\mathcal{N}(\sigma_d) = \sigma_{rd}$  (up to a change in the ambient dimension, which is irrelevant by Lemma 5). Then (35) follows from the non-increase of trace distance under quantum operations.  $\square$

(It seems plausible that Lemma 8 should also hold for non-integer  $r$  whenever  $rd_1$  and  $rd_2$  are integers, but we know of no proof of this claim.)

Now we turn to the proof of Theorem 3.

*Proof of Theorem 3.* The case when  $k \leq d$  is the simple one. Using first the triangle inequality, and then Lemma 6, we find

$$\begin{aligned} & \|\text{Schur}(k, d) - \text{Schur}(k, d/r)\|_1 \\ & \leq \|\text{Schur}(k, d) - \text{Planch}(k)\|_1 + \|\text{Schur}(k, d/r) - \text{Planch}(k)\|_1 \end{aligned} \quad (37)$$

$$\leq \sqrt{2} \frac{k + kr}{d} \leq O(kr/d). \quad (38)$$

This establishes the  $\Omega(d/r)$  lower bound on the number of copies.

Now, for any choice of  $d$  and  $r$ , we would like to show that  $k = d/r$  copies is sufficient to obtain  $\|\text{Schur}(k, d) - \text{Schur}(k, d/r)\|_1 \geq \Omega(1)$ . The achievability part of Theorem 3 will then follow from the same sort of amplification argument we used in the proof of Corollary 7. Choose  $\nu$  to be the smallest positive integer satisfying  $r^\nu \geq 72\sqrt{2}$ ; note that  $\nu \leq 7$ . Then we use both the upper and lower bounds of Lemma 6 to find

$$\left\| \text{Schur}\left(\frac{d}{r}, \frac{d}{r}\right) - \text{Planch}\left(\frac{d}{r}\right) \right\|_1 \geq \frac{1}{36} \quad (39)$$

$$\left\| \text{Schur}\left(\frac{d}{r}, dr^{\nu-1}\right) - \text{Planch}\left(\frac{d}{r}\right) \right\|_1 \leq \frac{\sqrt{2}}{r^\nu} \leq \frac{1}{72}. \quad (40)$$

Next we abbreviate  $S_{k,d} := \text{Schur}(k, d)$  and  $P_k := \text{Planch}(k)$  and use the triangle inequality to bound

$$\begin{aligned} & \left\| S_{\frac{d}{r}, \frac{d}{r}} - S_{\frac{d}{r}, d} \right\|_1 + \left\| S_{\frac{d}{r}, d} - S_{\frac{d}{r}, dr} \right\|_1 + \cdots + \left\| S_{\frac{d}{r}, dr^{\nu-2}} - S_{\frac{d}{r}, dr^{\nu-1}} \right\|_1 \\ & \geq \left\| S_{\frac{d}{r}, \frac{d}{r}} - P_{\frac{d}{r}} \right\|_1 - \left\| S_{\frac{d}{r}, dr^{\nu-1}} - P_{\frac{d}{r}} \right\|_1 \geq \frac{1}{72}. \end{aligned} \quad (41)$$

According to Lemma 8, the first term on the left hand side of (41) is the largest. Thus,

$$\left\| \text{Schur}\left(\frac{d}{r}, \frac{d}{r}\right) - \text{Schur}\left(\frac{d}{r}, d\right) \right\|_1 \geq \frac{1}{72\nu} \geq \frac{1}{504}. \quad (42)$$

This concludes the proof of Theorem 3.  $\square$

Our only remaining task is to prove Lemma 6, establishing tight bounds on the distance between  $\text{Schur}(k, d)$  and  $\text{Planch}(k)$ . Here our primary tools are explicit formulas for the Plancherel averages of certain functions of Young diagrams over  $\mathcal{S}_k$ .

*Proof of Lemma 6.* Let  $\Delta_{k,d}$  denote the total variation distance between  $\text{Schur}(k, d)$  and  $\text{Planch}(k)$ . According to (26) and (27),  $\Delta_{k,d} := \|\text{Schur}(k, d) - \text{Planch}(k)\|_1$  can be written

$$\Delta_{k,d} = \mathbf{E}_{\lambda \vdash k} \left| \prod_{(i,j) \in \lambda} \left( 1 + \frac{j-i}{d} \right) - 1 \right| \quad (43)$$

where the expectation is taken over  $\text{Planch}(k)$ .

We will now derive both upper and lower bounds on  $\Delta_{k,d}$  from (43), starting with the  $O(k/d)$  upper bound. By the Cauchy-Schwarz inequality (or concavity of the square root and Jensen's inequality),

$$\Delta_{k,d}^2 \leq \mathbf{E}_{\lambda \vdash k} \left[ \prod_{(i,j) \in \lambda} \left( 1 + \frac{j-i}{d} \right) - 1 \right]^2 \quad (44)$$

$$= \mathbf{E}_{\lambda \vdash k} \left[ \prod_{(i,j) \in \lambda} \left( 1 + \frac{j-i}{d} \right)^2 - 2 \prod_{(i,j) \in \lambda} \left( 1 + \frac{j-i}{d} \right) + 1 \right] \quad (45)$$

$$= \mathbf{E}_{\lambda \vdash k} \prod_{(i,j) \in \lambda} \left( 1 + \frac{j-i}{d} \right)^2 - 1, \quad (46)$$

where in the last line we have used linearity of expectation and the fact that  $\text{Schur}(k, d)$  is normalized. Now using the inequality  $1 + x \leq e^x$ , we have

$$\Delta_{k,d}^2 \leq \mathbf{E}_{\lambda \vdash k} \exp \left( 2 \sum_{(i,j) \in \lambda} \frac{j-i}{d} \right) - 1 \quad (47)$$

$$= \sum_{m=1}^{\infty} \frac{2^m}{m! d^m} \mathbf{E}_{\lambda \vdash k} v_1(\lambda)^m \quad (48)$$

where

$$v_1(\lambda) := \sum_{(i,j) \in \lambda} (j-i) \quad (49)$$

can be viewed as a measure of the asymmetry of  $\lambda$ .

Intuitively, we expect this to be small since typical Plancherel-distributed Young diagrams are close to a certain limiting shape that is symmetric, and in which the length of the longest row,  $\lambda_1$ , is close to  $2\sqrt{k}$  [38, 52] (in fact,  $\lambda_1$  is exponentially tightly concentrated around this value [7]). Our proof will use results obtained by Kerov in the course of describing the asymptotically Gaussian fluctuations about the limiting shape [35].

It turns out that  $v_1$  can be expressed exactly in terms of a central character of the symmetric group [34], namely

$$v_1(\lambda) = \binom{k}{2} \frac{\chi_{\lambda}(\tau)}{\dim \mathcal{P}_{\lambda}} \quad (50)$$

where  $\tau$  is any representative of the conjugacy class of transpositions. Then using the results of Kerov [35, equation (2.5) and Lemma 2.3], we find

$$\mathbf{E}_{\lambda \vdash k} v_1(\lambda)^{2m-1} = 0 \quad (51)$$

$$\mathbf{E}_{\lambda \vdash k} v_1(\lambda)^{2m} = \frac{(2m)! k!}{4^m m! (k-2m)!} \leq \frac{(2m)! k^{2m}}{4^m m!} \quad (52)$$

for  $m = 1, 2, \dots$ . Using these expressions in (48), we find

$$\Delta_{k,d}^2 \leq \sum_{m=1}^{\infty} \frac{1}{m!} \left( \frac{k}{d} \right)^{2m} = e^{(k/d)^2} - 1 \leq 2 \left( \frac{k}{d} \right)^2, \quad (53)$$

where in the second inequality we have assumed  $k \leq d$ . This shows that  $\Delta_{k,d} = o(1)$  provided  $k = o(d)$ , and indeed, that  $\Delta_{k,d}$  is exponentially small (in  $\log d$ ) if  $k = d^{1-\epsilon}$  for any fixed  $\epsilon > 0$ .

Now we establish the  $\Omega(k/d)$  lower bound for  $\|\text{Planch}(k) - \text{Schur}(k, d)\|_1$ , again by a careful examination of the distribution of  $v_1(\lambda)$ . If  $S$  is any subset of the partitions of  $k$ , then

$$\Delta_{k,d} \geq \mathbf{E}_{\lambda \vdash k} \left[ 1 - \prod_{(i,j) \in \lambda} \left( 1 + \frac{j-i}{d} \right) \right] \delta(\lambda \in S) \quad (54)$$

$$= \Pr(\lambda \in S) - \mathbf{E}_{\lambda \vdash k} \prod_{(i,j) \in \lambda} \left( 1 + \frac{j-i}{d} \right) \delta(\lambda \in S) \quad (55)$$

$$\geq \Pr(\lambda \in S) - \mathbf{E}_{\lambda \vdash k} e^{v_1(\lambda)/d} \delta(\lambda \in S) \quad (56)$$

$$\geq \Pr(\lambda \in S) \left( 1 - \max_{\lambda \in S} e^{v_1(\lambda)/d} \right). \quad (57)$$

Thus, if we can show that  $v_1(\lambda)$  has a reasonable probability of being very negative, we will find that  $\Delta_{k,d}$  is appreciable. Under the Plancherel distribution, the probability of a diagram  $\lambda$  and its transposed diagram  $\lambda^T$  (satisfying  $v_1(\lambda^T) = -v_1(\lambda)$ ) are equal, so by symmetry,

$$\Pr(v_1(\lambda) \leq -v) = \frac{1}{2} \Pr(v_1(\lambda)^2 \geq v^2). \quad (58)$$

Letting  $S$  be the set of  $\lambda$ 's with  $v_1(\lambda) \leq -v$  (for some  $v > 0$  to be specified later), we have

$$\Delta_{k,d} \geq \frac{1}{2} \Pr(v_1(\lambda)^2 \geq v^2)(1 - e^{-v/d}). \quad (59)$$

Now we use (52) for  $m = 1$  and  $m = 2$ , i.e.,

$$\mathbf{E}_{\lambda \vdash k} v_1(\lambda)^2 = \binom{k}{2} \quad (60)$$

$$\mathbf{E}_{\lambda \vdash k} v_1(\lambda)^4 = \frac{3}{4} k(k-1)(k-2)(k-3), \quad (61)$$

to show that

$$\mathbf{E}_{\lambda \vdash k} v_1(\lambda)^4 - \left[ \mathbf{E}_{\lambda \vdash k} v_1(\lambda)^2 \right]^2 = \frac{1}{4} k(k-1)(2k^2 - 15k + 18) \leq \frac{k^4}{2}. \quad (62)$$

Applying a one-sided variant of the Chebyshev inequality, namely [47]

$$\Pr(X \geq \mu - \alpha\sigma) \geq \frac{\alpha^2}{1 + \alpha^2} \quad (63)$$

where  $\mu := \mathbf{E} X$  and  $\sigma^2 := \mathbf{E} X^2 - \mu^2$ , we find

$$\Pr\left(v_1(\lambda)^2 \geq \binom{k}{2} - \alpha \frac{k^2}{\sqrt{2}}\right) \geq \frac{\alpha^2}{1 + \alpha^2}. \quad (64)$$

Choosing (say)  $\alpha = \frac{1}{2\sqrt{2}}$ , we have  $\Pr(v_1(\lambda)^2 \geq \frac{1}{4}k^2 - \frac{1}{2}k) \geq \frac{1}{9}$ ; plugging this into (59) gives

$$\Delta_{k,d} \geq \frac{1}{18} \left[ 1 - \exp\left(-\frac{k}{2d} \sqrt{1 - \frac{2}{k}}\right) \right], \quad (65)$$

which is  $\Omega(1)$  provided  $k = O(d)$ . To obtain a specific numeric bound, we divide into cases. If  $2 \leq k \leq 5$ , then  $\Delta_{k,d} \geq \Delta_{2,d} = 1/d \geq k/36d$ , where the equality can be verified by substituting  $k = 2$  into (43). If  $6 \leq k \leq d$ , then  $\frac{1}{2}(1 - \frac{2}{k})^{1/2} \geq \frac{1}{\sqrt{6}}$ , so  $\Delta_{k,d} \geq \frac{1}{18}(1 - e^{-k/\sqrt{6}d})$ . Next, the convexity of  $e^{-x/\sqrt{6}}$  implies that  $e^{-x/\sqrt{6}} \geq 1 - xe^{-1/\sqrt{6}} \geq 1 - x/2$ . Thus  $\Delta_{k,d} \geq k/36d$ .  $\square$

To put the results of this section in context, we review the situation when  $k \gg d$  or  $k \ll d$ . When  $k \rightarrow \infty$  with  $d$  held constant, applying the measurement  $\{\Pi_\lambda\}_{\lambda \vdash k}$  to  $\rho^{\otimes k}$  and outputting  $\bar{\lambda} := \lambda/k$  has long been known to be a valid estimator of the spectrum of  $\rho$  [36]. Indeed, if  $r_1 \geq \dots \geq r_d$  are the eigenvalues of  $\rho$ , then [16, 32] proved that

$$\text{tr } \Pi_\lambda \rho^k \leq (k+1)^{d(d-1)/2} \exp(-kD(\bar{\lambda}||r)), \quad (66)$$

where  $D(p||q) := \sum_i p_i \log(p_i/q_i)$  is the (classical) relative entropy. In fact, this coefficient of  $k$  in the exponential is optimal as  $k \rightarrow \infty$ . This inequality is usually only interesting when  $k = \Omega(d^2)$ , so our Theorem 3 can be viewed as the first positive result for spectrum estimation in the regime where  $k = o(d^2)$ .

What about the other limit, when  $k \ll d$ ? The only constructive result previously known was the swap test, which yields  $\Delta_{2,d} = 1/d$ . However,

$$F(\text{Schur}(2, d), \text{Planch}(2)) = \frac{1}{2} \left( \sqrt{1 - \frac{1}{d}} + \sqrt{1 + \frac{1}{d}} \right) = 1 - O(1/d^2), \quad (67)$$

so repeating swap tests will actually require  $k = \Omega(d^2)$  copies before achieving  $\Omega(1)$  distinguishability. Indeed,  $k = \Omega(d^2)$  copies are needed for  $\Omega(1)$  distinguishability even if we perform the optimal individual measurement on  $O(1)$  copies at a time. This follows from the following variant of Lemma 6:

**Lemma 9.**

$$F(\text{Schur}(k, d), \text{Planch}(k)) \geq 1 - \frac{k^3}{12d^2} \quad (68)$$

$$F(\text{Schur}(k, d), \text{Planch}(k)) \geq 1 - C^{-\sqrt{k}} - \frac{9k^2}{4d^2}, \quad (69)$$

where  $C > 1$  is a universal constant.

These bounds are not quite strong enough to match the  $1 - O(k^2/d^2)$  upper bound on the fidelity from Lemma 6 (which we conjecture is tight), but they still prove that achieving the performance of the optimal measurement requires operations that are entangled across many copies.

*Proof.* We begin with a direct calculation of

$$f_{k,d} := F(\text{Schur}(k, d), \text{Planch}(k)) = \mathbf{E}_{\lambda \vdash k} \sqrt{\prod_{(i,j) \in \lambda} \left(1 + \frac{j-i}{d}\right)}. \quad (70)$$

Consider the contribution from a particular  $\lambda$ , and suppose for now that  $\lambda = \lambda^T$ . Then  $(i, j) \in \lambda$  if and only if  $(j, i) \in \lambda$ , so each factor of  $1 + (j-i)/d$  is accompanied by a corresponding factor of  $1 - (j-i)/d$  (unless  $i = j$ , in which case this term is 1 and can be ignored). Combining these terms, the entire product can be written as

$$\sqrt{\prod_{(i,j) \in \lambda} \left(1 + \frac{j-i}{d}\right)} = \left[ \prod_{(i,j) \in \lambda} \left(1 - \frac{(j-i)^2}{d^2}\right) \right]^{\frac{1}{4}}. \quad (71)$$

Otherwise, suppose that  $\lambda \neq \lambda^T$ . Recall that  $\lambda$  and  $\lambda^T$  have the same probability under the Plancherel distribution. Thus the terms corresponding to  $\lambda$  and  $\lambda^T$  appear in (70) with the same coefficient. Next we use the identity  $x + y \geq 2\sqrt{xy}$  (for  $x, y \geq 0$ ) to show

$$\sqrt{\prod_{(i,j) \in \lambda} \left(1 + \frac{j-i}{d}\right)} + \sqrt{\prod_{(i,j) \in \lambda^T} \left(1 + \frac{j-i}{d}\right)} \geq 2 \left[ \prod_{(i,j) \in \lambda} \left(1 - \frac{(j-i)^2}{d^2}\right) \right]^{\frac{1}{4}}. \quad (72)$$

Combining (71) and (72), we can lower bound (70) as

$$f_{k,d} \geq \mathbf{E}_{\lambda \vdash k} \left[ \prod_{(i,j) \in \lambda} \left(1 - \frac{(j-i)^2}{d^2}\right) \right]^{\frac{1}{4}} \geq 1 - \frac{k^3}{12d^2}, \quad (73)$$

where the last inequality uses the fact that the product is minimized for the diagram  $\lambda = (k)$ . This proves (68).

The proof of (69) is mostly the same, but also uses bounds on large deviations of  $\lambda_1$  for Plancherel-distributed  $\lambda$ . In particular, equation (1.20) of [7] implies that under  $\text{Planch}(k)$ ,

$$\Pr(\lambda_1 \geq 3\sqrt{k}) \leq C^{-\sqrt{k}} \quad (74)$$

for a universal constant  $C > 1$  (in fact,  $C = \exp(6 \cosh^{-1}(3/2) + 2\sqrt{5}) \approx 28189$ ). Since  $(i, j) \in \lambda$  implies that  $j - i \leq \lambda_1 - 1$ , we can lower bound the middle expression in (73) to give

$$f_{k,d} \geq \Pr(\lambda_1 \leq 3\sqrt{k}) \left(1 - \frac{9k}{d^2}\right)^{\frac{k}{4}} \geq 1 - C^{-\sqrt{k}} - \frac{9k^2}{4d^2}, \quad (75)$$

which completes the proof.  $\square$

It is interesting to consider the classical analog of applying the swap test to the collision problem. If we compare  $k/2$  sets of 2 objects each, then estimating the number of collisions up to a constant multiple requires  $k = \Omega(d)$ . This is also quadratically worse than the optimal collective estimation procedure on  $k$  objects, but for apparently different reasons.

## 4.2 A black box for the quantum collision problem

We now define a version of this problem that allows for adaptive algorithms by introducing a quantum oracle. The oracle implements the isometry

$$|i\rangle \mapsto |i\rangle |\psi_{f(i)}\rangle, \quad (76)$$

where  $\mathcal{B} := \{|\psi_1\rangle, \dots, |\psi_d\rangle\}$  is an arbitrary (unknown) orthonormal basis of  $\mathbb{C}^d$  and  $f$  is either a one-to-one function or an  $r$ -to-one function. The goal is to determine which is the case using as few queries as possible. In fact, we assume that the isometry is extended to a unitary operator  $R$  acting on  $\mathbb{C}^d \otimes \mathbb{C}^d$  by

$$|i\rangle |y\rangle \mapsto |i\rangle U|y \oplus f(i)\rangle, \quad (77)$$

where  $U := \sum_i |\psi_i\rangle\langle i|$  is the unitary matrix effecting a transformation from the standard basis to  $\mathcal{B}$ . We also assume we can perform its inverse  $R^\dagger$ .

By considering the case where the basis  $\mathcal{B}$  (or equivalently  $U$ ) is known, it is clear that the quantum lower bound for the usual collision problem [1] implies an  $\Omega(\sqrt[3]{d/r})$  lower bound on the number of queries for the quantum collision problem as well. We present an algorithm for this problem that requires only  $O(\sqrt[3]{d/r} \log d/r)$  queries, nearly matching the lower bound.

The basic tool used in this algorithm is the following amplified version of the swap test:

**Lemma 10 (Amplified swap test with small disturbance).** *Suppose we are given a quantum state  $\rho$  satisfying*

$$\frac{1}{2} \|\rho - |\psi\rangle\langle\psi|\|_1 \leq \epsilon \quad (78)$$

where

$$|\psi\rangle = \sum_i a_i |\alpha_i\rangle^{\otimes m} |\beta_i\rangle^{\otimes m} |\gamma_i\rangle \quad (79)$$

with  $\langle\gamma_i|\gamma_j\rangle = \delta_{i,j}$  and  $\sum_i |a_i|^2 = 1$ . We further assume that for each  $i$ ,  $\theta_i := \langle\alpha_i|\beta_i\rangle$  is equal to either 0 or 1. Then there exists a unitary operation  $U$ , implementable with  $O(m)$  gates, such that

$$\frac{1}{2} \|U\rho U^\dagger - |\psi'\rangle\langle\psi'|\|_1 \leq \epsilon + 2^{1-m/2} \quad (80)$$

where

$$|\psi'\rangle := \sum_i a_i (-1)^{\theta_i} |\alpha_i\rangle^{\otimes m} |\beta_i\rangle^{\otimes m} |\gamma_i\rangle. \quad (81)$$

*Proof.* The procedure for implementing  $U$  is as follows. First perform the  $m$ -fold swap test without the final measurements. Conditioned on any swap test showing that the states are orthogonal, record this answer in an ancilla qubit; otherwise, record that they are the same. Apply a phase of  $-1$  in the “same” case. Finally, uncompute the ancilla qubit along with all the swap tests.

To analyze  $U$ , we initially assume we are acting on  $|\psi\rangle\langle\psi|$ . Let  $\Pi$  denote the projector that results from each swap test returning the symmetric outcome, so that  $U = 1 - 2\Pi$ . Since the swap test will always project  $|\alpha_i\rangle|\alpha_i\rangle$  onto the symmetric space, we have

$$\Pi|\alpha_i\rangle^{\otimes m} |\beta_i\rangle^{\otimes m} |\gamma_i\rangle = |\alpha_i\rangle^{\otimes m} |\beta_i\rangle^{\otimes m} |\gamma_i\rangle \quad (82)$$

whenever  $\theta_i = 1$ . On the other hand, if  $\theta_i = 0$  then

$$\Pi|\alpha_i\rangle^{\otimes m} |\beta_i\rangle^{\otimes m} |\gamma_i\rangle = \frac{1}{\sqrt{2^m}} |\eta_i\rangle^{\otimes m} |\gamma_i\rangle \quad (83)$$

where  $|\eta_i\rangle := \frac{1}{\sqrt{2}} (|\alpha_i\rangle|\beta_i\rangle + |\beta_i\rangle|\alpha_i\rangle)$ . Combining these, we obtain

$$U|\psi\rangle = \sum_i a_i ((-1)^{\theta_i} |\alpha_i\rangle^{\otimes m} |\beta_i\rangle^{\otimes m} - \delta_{\theta_i,0} 2^{1-\frac{m}{2}} |\eta_i\rangle^{\otimes m}) |\gamma_i\rangle \quad (84)$$

$$= |\psi'\rangle - \sum_i a_i \delta_{\theta_i,0} 2^{1-\frac{m}{2}} |\eta_i\rangle^{\otimes m} |\gamma_i\rangle. \quad (85)$$

Now the fidelity between  $U|\psi\rangle$  and  $|\psi'\rangle$  is

$$F(|\psi'\rangle\langle\psi'|, U|\psi\rangle\langle\psi|U^\dagger) = |\langle\psi'|U|\psi\rangle|^2 = \left|1 - \sum_i a_i \delta_{\theta_i, 0} 2^{1-\frac{m}{2}} \langle\psi'|\eta_i\rangle^{\otimes m} |\gamma_i\rangle\right|^2 \quad (86)$$

$$= \left|1 - \sum_i |a_i|^2 \delta_{\theta_i, 0} 2^{1-m}\right|^2 \geq |1 - 2^{1-m}|^2 \geq 1 - 2^{2-m}. \quad (87)$$

Finally,

$$\frac{1}{2} \|U\rho U^\dagger - |\psi'\rangle\langle\psi'|\|_1 \leq \frac{1}{2} \|U\rho U^\dagger - U|\psi\rangle\langle\psi|U^\dagger\|_1 + \frac{1}{2} \|U|\psi\rangle\langle\psi|U^\dagger - |\psi'\rangle\langle\psi'|\|_1 \quad (88)$$

$$\leq \frac{1}{2} \|\rho - |\psi\rangle\langle\psi|\|_1 + 2^{1-\frac{m}{2}} \leq \epsilon + 2^{1-\frac{m}{2}}, \quad (89)$$

where (88) uses the triangle inequality and (89) uses first the unitary invariance of the trace norm and then (87) and (30).  $\square$

We will use this amplified swap test to prove

**Theorem 11.** *The quantum  $r$ -collision problem has query complexity  $O(\sqrt[3]{d/r} \log d/r)$ .*

*Proof.* We give an explicit algorithm based on the quantum algorithm for the usual collision problem [11]. We first outline that original algorithm and then show how to extend it to the quantum problem.

The algorithm builds a table of the first  $\sqrt[3]{d/r}$  items and uses Grover's algorithm to search the remaining items for an entry of the table. Set  $K := \{1, \dots, \sqrt[3]{d/r}\}$  and  $\bar{K} := \{1, \dots, d\} - K$ . Then the algorithm consists of two steps:

1. Build the table  $L := \{(i, f(i)) : i \in K\}$ .
2. Use Grover search to find  $j \in \bar{K}$  such that  $F(j) = 1$ , where the function  $F : \bar{K} \rightarrow \{0, 1\}$  is defined by  $F(j) = 1$  if there is  $i \in K$  with  $f(i) = f(j)$  and  $F(j) = 0$  otherwise.

The first step uses  $O(\sqrt[3]{d/r})$  queries. In the second step, the function  $F$  must be computed reversibly, so two queries ( $R$  and  $R^\dagger$ ) are used for each call to  $F$ . Note that it is important that  $F$  can be computed on a superposition of inputs. The total number of queries in the second step is twice the number of Grover iterations needed to find a marked item, which we now analyze.

Assume that  $f$  is an  $r$ -to-one function. The probability that there is any collision among the elements of  $K$  is only  $O((r/d)^{1/6})$ , implying that the number of  $j \in \bar{K}$  with  $F(j) = 1$  is  $(r-1)(d/r)^{1/3}$  with high probability. Therefore, the number of Grover iterations necessary to find a solution is  $O(\sqrt{d/[r(d/r)^{1/3}]}) = O((d/r)^{1/3})$ . In summary, the total number of queries of the above algorithm is  $O((d/r)^{1/3})$ .

Now we adapt this algorithm to the quantum collision problem. Some adaptation is necessary because it is not possible to directly check equality of two quantum states  $|\psi_j\rangle$  and  $|\psi_{j'}\rangle$ . However, using Lemma 10, we can effectively test equality using  $m := 2 + 2 \log d/r$  copies of the quantum states, increasing the query complexity only by a factor of  $O(\log d/r)$ .

For this to work, it is important that we can reuse the states corresponding to the entries in the table. Iterating Lemma 10, we find that the error after  $\ell$  Grover iterations is at most  $\ell \cdot 2^{1-m/2} \leq \ell r/d$ . Since the number of Grover iterations is  $\ell = O((d/r)^{1/3})$ , the total error is asymptotically negligible, and we obtain essentially the same performance as in the classical collision problem.  $\square$

So far, we have only considered the query complexity of the quantum collision problem, without worrying about the algorithm's running time. The quantum algorithm for the classical collision problem can run in time  $O(\sqrt[3]{d/r} \log d/r)$  (i.e., only logarithmically greater than the query complexity), since after sorting the table  $L$  (in time  $O(\sqrt[3]{d/r} \log d/r)$ ), the function  $F$  can



be computed in time  $O(\log d/r)$  [11]. For the quantum collision problem, there is no obvious way to preprocess the table of quantum states. However, by using Grover's algorithm to implement  $F$  in time  $O(\sqrt{(d/r)^{1/3}} \log d/r)$ , we can obtain a running time of  $O(\sqrt{d/r} \log d/r)$ —substantially greater than the query complexity, but better than implementing  $F$  by brute force search. Note that to achieve these running times (for the both classical and quantum collision problems), we must assume a quantum RAM model. In other words, we must assume that it is possible to access any desired location in the table—or any superposition thereof—in constant time (see for example [42, Section 6.5]). If we restrict ourselves to the quantum circuit model, not allowing parallelization, then the running time is even longer for both problems.

We conclude this section by briefly considering some alternate versions of the quantum collision problem, one that has a fairly simple solution, and two about which less is known.

- *Input in an unknown basis:* Our definition of the black box for the quantum collision problem in (77) gives the output in an unknown basis. Alternatively, we can consider a problem where the input is taken in an unknown basis. Here we define the oracle  $R$  to act as  $R[U|i\rangle \otimes |y\rangle] = U|i\rangle \otimes |y \oplus f(i)\rangle$ , where  $U$  is again an unknown unitary and  $f$  is either one-to-one or  $r$ -to-one, and the problem is to determine which is the case. This problem was suggested by Scott Aaronson; the solution we give here was also found independently by Andris Ambainis.

The algorithm first queries  $R$  on  $(d/r)^{1/3}$  basis states of the form  $|j\rangle \otimes |0\rangle$  for uniformly random  $j$ , and then measures the second register, producing a table  $L$  of uniformly random function values  $f(i)$ , which are all distinct with high probability. Then we perform a Grover search for a state  $U|i\rangle$  such that  $f(i) \in L$ . There are two components to the Grover search. First, we need to identify our target space. Let  $\Pi := \sum_{i \in f^{-1}(L)} U|i\rangle\langle i|U^\dagger$ . Note that  $\text{tr } \Pi$  is either  $(d/r)^{1/3}$  or  $r(d/r)^{1/3}$ , and that we can perform the reflection  $1 - 2\Pi$  using only two queries (one of  $R$  and one of  $R^\dagger$ ). This implementation can be done by brute force in time  $O(r(d/r)^{1/3})$ , or by Grover search in time  $O(\sqrt{r(d/r)^{1/3}})$ —either way, regardless of whether  $f$  is one-to-one or  $r$ -to-one, and using only two queries. Next, we need a starting state  $|\psi\rangle$  that we can efficiently map to and from the  $|0\rangle$  state, so that we can both initialize it and perform the reflection  $1 - 2|\psi\rangle\langle\psi|$ . One such  $|\psi\rangle$  is a state that is maximally entangled between the first register and an ancilla register of equal dimension. Then we have  $\langle\psi|(I \otimes \Pi)|\psi\rangle = \text{tr } \Pi/d$ .

The algorithm now alternates  $\ell$  times between the operations  $1 - 2|\psi\rangle\langle\psi|$  and  $1 - 2\Pi$ . We choose  $\ell$  so that if  $f$  is  $r$ -to-one then the probability of finding a solution is nearly 1, while if  $f$  is one-to-one then the probability is far from one. Thus, we will be able to distinguish the two cases with constant advantage. Following [12], we set

$$\ell := \left\lceil 4/\pi \sqrt{\sin^{-1}\left((r/d)^{\frac{2}{3}}\right)} \right\rceil = \Theta((d/r)^{1/3}). \quad (90)$$

The probability of finding a solution is  $\geq 1 - O(r^{1/3}/d^{2/3})$  if  $f$  is  $r$ -to-one and is  $\leq 1/r + O(r^{1/3}/d^{2/3})$  if  $f$  is one-to-one.

If instead both the input and output bases are unknown, then we know of no algorithm better than making  $O(d)$  queries with random inputs and using the algorithm of Section 4.1.

- *No access to  $R^\dagger$ :* Suppose that instead of the full unitary oracle (77),  $R$  is simply the isometry (76), mapping  $|i\rangle$  to  $|i\rangle|\psi_{f(i)}\rangle$ . In this case, the best algorithm we know queries the oracle  $O(\log d)$  times on each of  $\sqrt{d}$  inputs (so that there is a collision among those inputs with reasonably high probability), and then searches the resulting table (either by brute force or by Grover search; in either case, making no further queries) using Lemma 10 to perform comparisons. We conjecture that this is optimal (up to the log factor), but we know of no lower bound better than the  $\Omega(d^{1/3})$  bound from the classical problem.
- *Black box rank estimation:* Suppose again that  $R$  is an isometry, but now mapping  $|i\rangle$  to  $|i\rangle|\psi_i\rangle$ , and we are only promised that either  $\{|\psi_i\rangle\}$  forms a basis for  $\mathbb{C}^d$  or that

$\frac{1}{d} \sum_i |\psi_i\rangle\langle\psi_i|$  is proportional to a projector onto a  $(d/r)$ -dimensional subspace. Here the best algorithm we know is to query  $R$  on  $O(d)$  inputs and use the rank estimation procedure from Section 4.1. We conjecture that this is optimal, but again have only the lower bound of  $\Omega(d^{1/3})$ .

## 5 Discussion

In this article, we have shown that weak Fourier-Schur sampling typically provides insufficient information to solve the hidden subgroup problem. Nevertheless, it remains possible that Schur duality could be a useful tool for the HSP. Just as weak Fourier sampling can always be used to refine the state space into smaller subspaces, even when it alone fails to solve the HSP, so we can use weak Fourier-Schur sampling to decompose the space even further. The Schur decomposition has the additional complication that the refined subspaces are no longer simply tensor products of single-copy subspaces, but this may actually be an advantage since entangled measurements are known to be necessary for some groups.

In principle, strong Fourier-Schur sampling is guaranteed to provide enough information to solve the HSP, simply because the hidden subgroup states are always distinguishable with  $k = \text{poly}(\log |G|)$  copies. However, it would be interesting to find a new efficient quantum algorithm for some HSP based on strong Fourier-Schur sampling. Perhaps a first step in this direction would be to analyze the performance of measurement in a random basis, as has been studied extensively in the case of weak Fourier sampling [24, 40, 43, 48].

Moving away from our original motivation of the hidden subgroup problem, the quantum collision problem may be of independent interest. As discussed in the latter part of Section 4.1, our results on the quantum collision sampling problem can be viewed as an exploration of spectrum estimation with  $k = o(d^2)$  copies, but much remains unknown about that regime. Many open problems also remain regarding the black box version of the quantum collision problem: Can the running time be reduced? What if both input and output registers are in unknown bases? Is the inverse oracle essential? And is black box rank estimation really no easier than rank estimation by sampling?

## Acknowledgments

We thank Scott Aaronson, Andris Ambainis, Masahito Hayashi, Keiji Matsumoto, Pranab Sen, and Umesh Vazirani for helpful discussions. We also thank Patrick Hayden for organizing a Bellairs Research Institute workshop on representation theory in quantum information, at which the seeds for this work were planted. This work was supported in part by the National Science Foundation under grant PHY-456720, by the Army Research Office under grant W9111NF-05-1-0294, and by the European Commission under Marie Curie grant ASTQIT.

## References

- [1] S. Aaronson and Y. Shi, *Quantum lower bounds for the collision and the element distinctness problems*, J. ACM **51** (2004), 595–605, quant-ph/0111102, quant-ph/0112086.
- [2] D. Bacon, A. M. Childs, and W. van Dam, *From optimal measurement to efficient quantum algorithms for the hidden subgroup problem over semidirect product groups*, Proc. 46th IEEE Symposium on Foundations of Computer Science, 2005, pp. 469–478.
- [3] ———, *Optimal measurements for the dihedral hidden subgroup problem*, to appear in Chicago Journal of Theoretical Computer Science, quant-ph/0501044.
- [4] D. Bacon, I. L. Chuang, and A. W. Harrow, *Efficient quantum circuits for Schur and Clebsch-Gordan transforms*, quant-ph/0407082.
- [5] ———, *The quantum Schur transform: I. Efficient qudit circuits*, To Appear in Proc. 18th ACM-SIAM Symposium on Discrete Algorithms, 2007.

- [6] E. Bagan, M. A Ballester, R. Muñoz-Tapia, and O. Romero-Isart, *Purity estimation with separable measurements*, Phys. Rev. Lett. **95** (2005), quant-ph/0509087, 110504.
- [7] J. Baik, P. Deift, and K. Johansson, *On the distribution of the length of the longest increasing subsequence of random permutations*, J. AMS **12** (1999), math.CO/9810105, 1119–1178.
- [8] A. Barenco, A. Berthiaume, D. Deutsch, A. Ekert, R. Jozsa, and C. Macchiavello, *Stabilisation of quantum computations by symmetrisation*, SIAM J. Comput. (1997), quant-ph/9604028, 1541–1557.
- [9] R. Beals, *Quantum computation of Fourier transforms over symmetric groups*, Proc. 29th ACM Symposium on Theory of Computing, ACM Press, New York, 1997, pp. 48–53.
- [10] R. Boneh and R. Lipton, *Quantum cryptanalysis of hidden linear functions*, Proc. Advances in Cryptology, Lecture Notes in Computer Science **963**, 1995, pp. 424–437.
- [11] G. Brassard, P. Høyer, and A. Tapp, *Quantum cryptanalysis of hash and claw-free functions*, Proc. 3rd Latin American Symposium on Theoretical Informatics, Lecture Notes in Computer Science **1380**, 1998, pp. 163–169.
- [12] G. Brassard, P. Høyer, M. Mosca, and A. Tapp, *Quantum amplitude amplification and estimation*, Quantum Computation & Information ( S. J. Lomonaco H. E. Brandt, eds.), Contemporary Mathematics Series Millenium Volume, vol. 305, AMS, 2002, pp. 53–74.
- [13] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, *Quantum fingerprinting*, Phys. Rev. Lett. **87** (2001), quant-ph/0102001, 167902.
- [14] A. M. Childs and P. Wocjan, *On the quantum hardness of solving isomorphism problems as non-abelian hidden shift problems*, quant-ph/0510185.
- [15] D. Coppersmith, *An approximate Fourier transform useful in quantum factoring*, Technical Report RC 19642, IBM Research Division, Yorktown Heights, NY, 1994, quant-ph/0201067.
- [16] M. Christandl and G. Mitchison, *The spectra of density operators and the Kronecker coefficients of the symmetric group*, Commun. Math. Phys. **261** (2006), quant-ph/0409016, 789–797.
- [17] M. Ettinger and P. Høyer, *A quantum observable for the graph isomorphism problem*, quant-ph/9901029.
- [18] ———, *On quantum algorithms for noncommutative hidden subgroups*, Advances in Applied Mathematics **25** (2000), quant-ph/9807029, 239–251.
- [19] M. Ettinger, P. Høyer, and E. Knill, *Hidden subgroup states are almost orthogonal*, quant-ph/9901034.
- [20] K. Friedl, G. Ivanyos, F. Magniez, M. Santha, and P. Sen, *Hidden translation and orbit coset in quantum computing*, Proc. 35th ACM Symposium on Theory of Computing, ACM Press, New York, 2003, pp. 1–9.
- [21] C. A. Fuchs and J. van de Graaf, *Cryptographic distinguishability measures for quantum mechanical states*, IEEE Trans. Inf. Theory **45** (1999), quant-ph/9712042, 1216–1227.
- [22] D. Gavinsky, *Quantum solution to the hidden subgroup problem for poly-near-Hamiltonian groups*, Quantum Information and Computation **4** (2004), 229–235.
- [23] R. Goodman and N. R. Wallach, *Representations and Invariants of the Classical Groups*, Cambridge University Press, Cambridge, 1998.
- [24] M. Grigni, L. Schulman, M. Vazirani, and U. Vazirani, *Quantum mechanical algorithms for the nonabelian hidden subgroup problem*, Combinatorica **24** (2004), 137–154.
- [25] L. K. Grover, *A fast quantum mechanical algorithm for database search*, Proc. 28th ACM Symposium on Theory of Computing, 1996, pp. 212–219.
- [26] S. Hallgren, *Polynomial-time quantum algorithms for Pell’s equation and the principal ideal problem*, Proc. 34th ACM Symposium on Theory of Computing, 2002, pp. 653–658.
- [27] ———, *Fast quantum algorithms for computing the unit group and class group of a number field*, Proc. 37th ACM Symposium on Theory of Computing, 2005, pp. 468–474.
- [28] L. Hales and S. Hallgren, *An improved quantum Fourier transform algorithm and applications*, Proc. 41st IEEE Symposium on Foundations of Computer Science, 2000, pp. 515–525.

- [29] S. Hallgren, C. Moore, M. Rötteler, A. Russell, and P. Sen, *Limitations of quantum coset states for graph isomorphism*, Proc. 38th ACM Symposium on Theory of Computing, 2006, pp. 604–617, quant-ph/0511148, quant-ph/0511149.
- [30] S. Hallgren, A. Russell, and A. Ta-Shma, *The hidden subgroup problem and quantum computation using group representations*, Proc. 32nd ACM Symposium on Theory of Computing, 2000, pp. 627–635.
- [31] A. W. Harrow, *Applications of coherent classical communication and the Schur transform to quantum information theory*, Ph.D. thesis, Massachusetts Institute of Technology, 2005.
- [32] M. Hayashi and K. Matsumoto, *Quantum universal variable-length source coding*, Phys. Rev. A **66** (2002), quant-ph/0202001, 022311.
- [33] G. Ivanyos, F. Magniez, and M. Santha, *Efficient quantum algorithms for some instances of the non-abelian hidden subgroup problem*, International Journal of Foundations of Computer Science **14** (2003), quant-ph/0102014, 723–739.
- [34] A. Jucys, *Symmetric polynomials and the center of the symmetric group ring*, Rep. Math. Phys. **5** (1974), 107–112.
- [35] S. Kerov, *Gaussian limit for the Plancherel measure of the symmetric group*, Comptes Rendus Acad. Sci. Paris, Sér. I **316** (1993), 303–308.
- [36] M. Keyl and R. F. Werner, *Estimating the spectrum of a density operator*, Phys. Rev. A **64** (2001), quant-ph/0102027, 052311.
- [37] G. Kuperberg, *A subexponential-time quantum algorithm for the dihedral hidden subgroup problem*, SIAM Journal on Computing **35** (2005), quant-ph/0302112, 170–188.
- [38] B. F. Logan and L. A. Shepp, *A variational problem for random Young tableaux*, Adv. Math. **26** (1977), 206–222.
- [39] C. Moore, D. N. Rockmore, and A. Russell, *Generic quantum Fourier transforms*, Proc. 15th ACM-SIAM Symposium on Discrete Algorithms, 2004, pp. 778–787.
- [40] C. Moore, D. N. Rockmore, A. Russell, and L. J. Schulman, *The hidden subgroup problem in affine groups: Basis selection in Fourier sampling*, Proc. 15th ACM-SIAM Symposium on Discrete Algorithms, SIAM, Philadelphia, 2004, pp. 1113–1122, quant-ph/0211124, extended version available at quant-ph/0503095.
- [41] C. Moore, A. Russell, and L. J. Schulman, *The symmetric group defies strong Fourier sampling*, Proc. 46th IEEE Symposium on Foundations of Computer Science, 2005, pp. 479–490.
- [42] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, 2000.
- [43] J. Radhakrishnan, M. Rötteler, and P. Sen, *On the power of random bases in Fourier sampling: Hidden subgroup problem in the Heisenberg group*, Proc. International Colloquium on Automata, Languages and Programming, Lecture Notes in Computer Science **3580**, 2005, pp. 1399–1411.
- [44] O. Regev, *Quantum computation and lattice problems*, Proc. 43rd Symposium on Foundations of Computer Science, IEEE, Los Alamitos, CA, 2002, pp. 520–529.
- [45] ———, *A subexponential time algorithm for the dihedral hidden subgroup problem with polynomial space*, quant-ph/0406151.
- [46] A. Schmidt and U. Vollmer, *Polynomial time quantum algorithm for the computation of the unit group of a number field*, Proc. 37th ACM Symposium on the Theory of Computing, 2005, pp. 475–480.
- [47] H. L. Selberg, *Über eine Ungleichungen zur Ergänzung des Tchebycheffschen Lemmas*, Skandinavisk Aktuarietidskrift **23** (1940), 121–125.
- [48] P. Sen, *Random measurement bases, quantum state distinction and applications to the hidden subgroup problem*, quant-ph/0512085.
- [49] J. P. Serre, *Linear Representations of Finite Groups*, Graduate Texts in Mathematics, vol. 42, Springer, New York, 1977.
- [50] P. W. Shor, *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM Journal on Computing **26** (1997), quant-ph/9508027, 1484–1509.

- [51] R. P. Stanley, *Theory and application of plane partitions*, Studies in Appl. Math. **1** (1971), 167–187 and 259–279.
- [52] A. M. Vershik and S. V. Kerov, *Asymptotic behavior of the Plancherel measure of the symmetric group and the limit form of Young tableaux*, Dokl. Akad. Nauk SSSR **233** (1977), 1024–1027.
- [53] ———, *Asymptotic behavior of the maximum and generic dimensions of irreducible representations of the symmetric group*, Funct. Anal. Appl. **19** (1989), 21–31, English translation of Funk. Anal. i Prolizhen **19** (1985), no. 1, 25–26.